QUEUE-AWARE PERFORMANCE OPTIMIZATION OF HETEROGENEOUS NETWORKS

by

Fancheng Kong

Dissertation submitted in fulfilment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

School of Computing & Communications
Faculty of Science
University of Technology Sydney
Sydney, Australia

August 24, 2018

ABSTRACT

To meet surging traffic demands, heterogeneous networks (HetNets) enable a more flexible, targeted and economical deployment of new infrastructure versus tower-mounted macro-only systems, which are very expensive to deploy and maintain. For a high network spectrum efficiency, load balancing across different tiers can be achieved by optimizing the association between users and base stations (BSs). To achieve a high energy efficiency, proper controls of BSs' activation (on/off status) and deployment density can significantly avoid unnecessary BS power consumption. However, some practical conditions are not considered in existing studies.

(i) Previous studies usually assumed that BSs were always busy transmitting packets to their associated users, which characterized a worst case of the performance metrics. In practice, one BS can either be busy or idle, depending on its queuing condition, in which case the performance metrics such as the packet delay should be further studied with queuing taken into account.

(ii) With the assumption of continuous BS transmission in previous literatures, the network power consumption linearly increases with the number of BSs only. Practically, the power consumption of a BS in the idle state is much lower than that in the busy state, the tuning of the network design parameters, for example, the bandwidth allocation and the BS deployment density, thus have a significant impact on the BS busy/idle status, which in turn affects the network energy efficiency.

(iii) Most of the previous studies focus on a uniform user distribution. In reality, users might not be evenly distributed and may form a cluster in certain

hot area. In such cases, the user association optimization in a per-tier fashion would result in a poor user experience in the overloaded areas, and a per-station association scheme is thus preferable.

To address the above considerations, the thesis focuses on the optimization of both the network spectrum efficiency and the network energy efficiency with practical assumptions of queuing and non-uniform user distribution, which is elaborated in the following.

1) Delay-optimal biased user association in HetNets. A thinned Poisson point process model to characterize the locations of BSs in the busy state, and an explicit expression of the average traffic intensity of each tier is obtained. On that basis, an optimization problem is formulated to minimize the lower bound of the network mean queuing delay by tuning the biasing factor of each tier, which is shown to be a convex problem. The simulation results demonstrate that the network queuing performance can be significantly improved by properly tuning the biasing factor. It is further shown that the network mean queuing delay might be improved at the cost of a deterioration of the network signal-to-interference ratio (SIR) coverage, which indicates a performance tradeoff between real-time and non-real-time traffic in HetNets.

2) Queue-aware optimal bandwidth allocation in HetNets. Based on the queuing analysis, a minimization problem of the network average power consumption and a maximization problem of the network SIR coverage are formulated, which are shown to be convex and concave with respect to the bandwidth allocation to each tier, respectively. By using an approximation of the average traffic intensity, closed-form solutions are obtained for both problems. Simulation results of a 2-tier HetNet demonstrate that the network average power consumption and the SIR coverage can be significantly improved by the optimal bandwidth allocation.

3) Queue-aware energy efficient base station density optimization in HetNets. By further using the approximation that BSs of a tier have the same SIR coverage, the cumulative distribution function (CDF) of the traffic intensity of each tier is obtained. On that basis, a minimization problem of the network average power consumption is studied by optimally tuning the activation ratio of micro BSs under the quality of serive (QoS) constraints of the network mean queuing delay and the network SIR coverage. Numerical results demonstrate that if the idle power coefficient is below a certain threshold, the optimal activation ratio should equal the one to minimize the network average power consumption per area. Otherwise, the optimal activation ratio should be obtained according to the QoS constraints. It is further shown that universal frequency reuse (UFR) outperforms spectrum partitioning (SP) in terms of both energy efficiency and SIR coverage in the considered scenario.

4) Optimal biased association scheme with non-uniform user distribution in HetNets. A practical scenario is studied where one cell is overloaded due to the cluster of users. By maximizing the mean user utility in the area of this overloaded cell and its neighboring cells, the optimal biasing factor can be obtained. It is found that in the scenario where the overloaded cell is fully surrounded by a macro cell, the optimal biasing factor logarithmically decreases with the user's intensity of the overloaded cell. Numerical results demonstrate that the mean user rate of the overloaded cell and the whole network can be significantly improved by properly tuning the biasing factor of the overloaded cell.

Key words: Heterogeneous network, Queuing, Non-uniform user distribution, Biasing factor, Bandwidth allocation, BS deployment density, Network mean queuing delay, Network average power consumption.

STATEMENT OF CANDIDATE

This thesis is the result of a research candidate conducted with another University as part of a collaborative Doctoral degree.

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree to any other university or institution other than University of Technology Sydney.

I also certify that the thesis is an original piece of research and it has been written by me.

In addition, I certify that all information sources and literature used are indicated in the thesis.

Production Note:
Signature removed prior to publication.
. . . . . .

Fancheng Kong

ACKNOWLEDGMENTS

I would like to express my profound gratitude to my principal supervisors, Professor Yingjie Jay Guo and Professor Hongbo Zhu, for their support, encouragement, supervision and guidance during this research work. Their vision, technical knowledge, moral support and continuous guidance helped me to complete my work successfully. They always stood with me whenever I encountered any difficulties during my postgraduate studies.

I also benefited from a number of people during my graduate studies. I am highly thankful to my co-supervisor, Associate Professor Xinghua Sun, for his research collaboration, guidance and valuable suggestions throughout this study. He has been and still is an inspiring personality for me. I should also like to thank Dr. Peiyuan Qin and Dr. Can Ding from Global Big Data Technology Centre for their guidance and support during this study. I owe my sincere thanks to Genie Tan for helping me with my enrollment to UTS and preparing everything upon my arrival.

I sincerely appreciate the financial support of the UTS President's scholarship and UTS top-up scholarship from Graduate Research School that made this study and my stay in Australia possible.

I express gratitude to my friends in Global Big Data Technology Centre. In particular, I would like to thank Dr. Shulin Chen, Dr. Haihan Sun and Dr. Yijiang Nan that helped make my stay enjoyable in Sydney

Finally, I would like to thank my mum Jie Wang and my dad Qinghai Kong who have always been proud of me and were always there for me. I could never ask for more. I should also like to thank my girlfriend Selena Liang, for always

supporting me when I was in tough times.

Contents

Abstract **ii**

Acknowledgments **vi**

Table of Contents **ix**

1 Introduction **1**
1.1 The Prevalence of HetNets 1
1.2 Key Technologies in Designing HetNets 4
1.2.1 Review of Spectrum Efficiency Optimization in HetNets 4
1.2.2 Review of Delay Optimization in HetNets 8
1.2.3 Review of Energy Efficiency Optimization in HetNets 10
1.3 Challenges and Motivations 12
1.4 Thesis Contributions and Structure 15
1.4.1 Thesis Contributions 15
1.4.2 Thesis Structure 17

2 Queuing Analysis **19**
2.1 Coupled Queue Problem 20
2.1.1 Queuing Model 20
2.1.2 Spatial-Temporal Correlation of the Queues 22
2.2 Methodology to Decouple the Correlation 24
2.2.1 Stochastic Geometry 24
2.2.2 Independent Thinning 28
2.3 Average Traffic Intensity 28
2.3.1 Orthogonal Spectrum Partitioning 28
2.3.2 Universal Frequency Reuse 31
2.3.3 Simulation Results 35
2.4 Conclusions 36

3 Queue-Aware Delay-Optimal Biased Association Optimization in HetNets **37**
3.1 Introduction 38
3.2 System Model 40

3.3 Queuing Delay Optimization . . . 41
3.3.1 Relation Between Average Traffic Intensity and Association Probability . . . 41
3.3.2 Queuing Delay Optimization . . . 44
3.4 Simulation Results . . . 52
3.5 Conclusion . . . 60

4 Queue-Aware Optimal Bandwidth Allocation in HetNets 63
4.1 Introduction . . . 64
4.2 System Model . . . 66
4.3 Network Average Power Consumption Minimization . . . 67
4.3.1 Problem Formulation . . . 67
4.3.2 Explicit Solution . . . 70
4.4 Network SIR Coverage maximization . . . 72
4.4.1 Problem Formulation . . . 72
4.4.2 Explicit Solution . . . 74
4.5 Simulation Results . . . 76
4.6 Conclusion . . . 82

5 Queue-Aware Energy Efficient BS Density Optimization in HetNets 83
5.1 Introduction . . . 84
5.1.1 Energy Efficiency Optimization . . . 84
5.1.2 QoS Constraint . . . 85
5.1.3 Universal Frequency Reuse Versus Spectrum Partitioning . . . 86
5.2 System Model . . . 88
5.3 Traffic Intensity Characterization . . . 88
5.3.1 Average Traffic Intensity . . . 89
5.3.2 Cumulative Distribution Function of Traffic Intensity . . . 93
5.4 QoS Constrained Network Average Power Consumption Optimization . . . 95
5.4.1 Performance Metrics . . . 96
5.4.2 QoS Constrained Network Average Power Consumption Minimization 100
5.5 Simulation Results . . . 100
5.5.1 Traffic Intensity . . . 101
5.5.2 Performance Metrics . . . 102
5.5.3 Optimal Activation Ratio . . . 105
5.6 Comparison of Universal Frequency Reuse and Spectrum Partitioning . . . 106
5.7 Conclusions . . . 108

6 Optimal Biased Association Scheme with Non-Uniform User Distribution 111
6.1 Introduction . . . 112
6.2 System Model . . . 114
6.2.1 Network Topology . . . 114

6.2.2 Association Region and Biasing Factor . . . 115
6.2.3 Mean User Rate . . . 117
6.3 Mean User Utility Optimization . . . 117
6.3.1 Probability of A Random User's Location . . . 118
6.3.2 Mean Logarithm of User Rate . . . 119
6.4 Case Study . . . 121
6.5 Simulation Results . . . 128
6.6 Conclusions . . . 134

7 Conclusions and Future Works **137**
7.1 Conclusions . . . 137
7.2 Future Works . . . 140

A Abbreviations **143**

B Publications **145**

Bibliography **147**